

Resource Accounting – Current Practices

Laura F. McGinnis

February, 2001

Justification

For GRID computing to be attractive to the entities who control the resources, the interface required by the GRID infrastructure must be as non-invasive as possible. “Joining the GRID”, i.e. making resources available to GRID users must require minimal changes to how labs and research centers currently manage their resources. GRID requirements should, therefore, be designed to sit on top of existing resource management processes, and to interface as seamlessly yet securely as possible. This design standard is being addressed in other GRID Working Groups; Distributed Accounting must also meet this challenge.

To build an effective interface between the GRID environment and existing resources, it is critical to understand how the existing resources are being managed. This paper is a compilation of survey results pertaining to allocation and usage accounting issues at various sites who have members participating in the Global GRID Forum. These sites (and others) were invited to describe their current allocation and usage accounting practices, to supply a baseline for defining that the GRID interface will need to work with.

Methodology

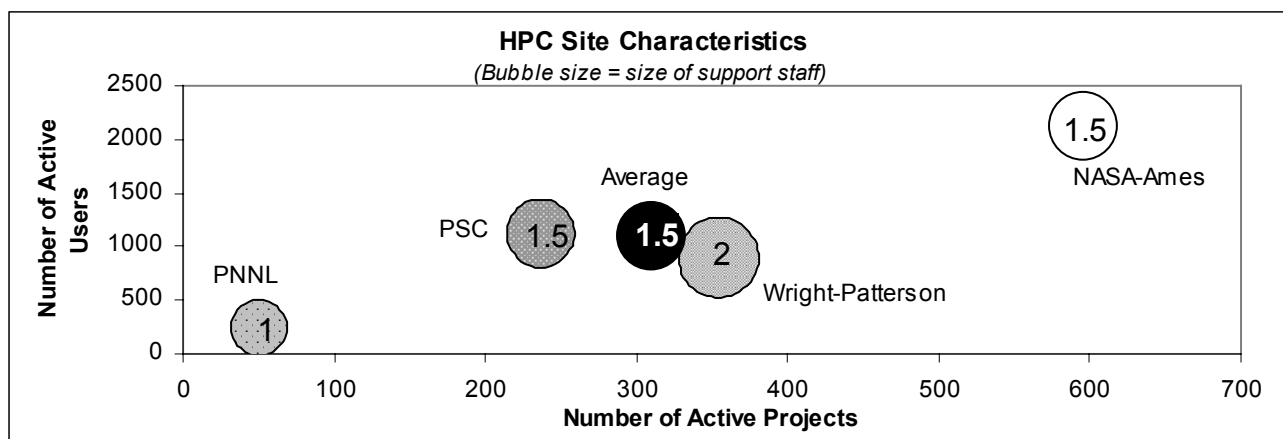
The survey was designed with open-ended questions which attempt to capture information relevant to account allocation and usage accounting and reporting. While the focus of the Distributed Accounting Working Group is usage accounting and reporting, it is clear that account allocation provides the framework for collecting and reporting resource usage data. Discussions within the Working Group also attempted to come up with definitions for the terms “account”, “accounting”, and “accountability”. Questions were included in the survey to gather input from sites on how these terms are commonly used.

Responses were collected via email¹ and are summarized in this report. A copy of the survey questions is given in Appendix A to this paper.

General Characteristics

- *How large are the sites?*
The 4 sites who responded to the survey average 1100 active users, distributed across an average of 310 different projects. PNNL is the smallest site, with 250 users working on about 50 projects; NASA-Ames reports the largest user population, at 2130 users, working on 596 projects.

¹ The 4 sites that responded to the survey are: Pacific Northwest National Laboratory (PNNL), Wright-Patterson AFB (ASC), NASA-Ames, and the Pittsburgh Supercomputing Center (PSC).



- *What are their staffing requirements for their current efforts?*
Each site has 1-2 people assigned to handle allocation and usage accounting tasks. Allocations are managed by an average of .8 FTE² and usage accounting required an average of .7 FTE. The tasks are generally split across more than one staff member, who have other responsibilities as well.
- *What resources are they managing?*
All sites are managing a variety of computing resources. Each site provides computing on platforms from a number of different vendors; i.e. no site is a single platform shop. As a result, each site already deals with diversity in system allocation and usage accounting processes, at the machine level.

NASA-Ames	
CRAY C90/Vector	16 processors, 8 GB
SGI Origin 2000/	24 processors, 7 GB
SGI Origin 2000/	256 processors, 64 GB
SGI Origin 2000/	64 processors, 16 GB
SGI Origin 2000/	512 processors, 196 GB
SGI Origin 2000/	8 processors, 2 GB
SGI Origin 2000	12 processors, 3 GB
PNNL	
P2SC IBM SP	512-node
P2SC IBM SP	8-node
Silver/NH1 IBM SP	20-node 96-processors
DELL Pentium Linux cluster	97-node 194-processors
PSC	
Compaq ES40 cluster	256 processors
Cray T3E	512 processors
Cray J90	8 processors
AlphaServer 8400 5/300	4 processors
Intel2 cluster	80 processors
CRAY J90 File Archiver	10 processors.
Wright-Patterson	
IBM SP Power 3	528 processors, 520 GB

² FTE = "Full-time Equivalent" employee; i.e. one person working a 40-hour work week.

IBM SP Power 2	256 processors, 254 GB
SGI Origin 2000	512 processors, 320 GB
Compaq GS-320	128 processors, 128 GB
Compaq ES-40	64 processors, 60 GB

Allocations

The way resource usage data is collected and reported is highly dependent on how access to those resources is allocated. The Distributed Accounting Working Group must consider allocation processes when defining standards for GRID accounting.

- *How are accounts established?*
Each of the sites has a unique process for determining allocations. The only point of commonality is the requirement that the request be reviewed prior to allocation. When and where the review occurs, its criteria, and what happens after a request is approved vary across all sites.
- *What are the review/approval requirements/criteria?*
Review criteria vary widely across the sites. Review criteria are defined by the mission of the facility. At Wright-Patterson, this means military significance; at PNNL and PSC, scientific merit. NASA-Ames requests are reviewed by NASA and forwarded to NASA-Ames for allocation; criteria include mission relevance (i.e. advancements in aeronautical research) as well as computational opportunities.

The duration of the review period also varies, even within a single site. At most sites, requests are reviewed quarterly or semiannually. At PSC, small grants (i.e. “starter” grants that allow users to develop code or become familiar with a resource’s capabilities) are allocated after internal peer review, which can happen on an ad hoc basis. Larger requests for PSC resources go through quarterly or semiannual review. At most sites³, industrial or commercial requests are generally not reviewed outside of the research facility.

- *How do users get access to resources?*
Generally, users are assigned to project groups, and the projects are allocated access to individual machines, or to a cluster of machines. All sites require the user to have a local account on the machine/resource to be used. At all sites, users access the resources that they have been allocated either interactively or by various batch management systems, although batch is the preferred access method. There are generally limits set of premiums charged for interactive access. Resource usage is charged against the allocation, i.e. at the project level.

Each site has a different policy for handling jobs that run over the project allocation:

- PNNL uses Qbank to keep track of current balances. When jobs are submitted, they include an estimate of how much of the

³ Wright-Patterson is the only site of the 4 that does not accept commercial accounts.

allocation the user expects to consume. If the remaining balance on the allocation is not sufficient, the job is rejected. If there are enough units available, a pending withdrawal is made on against the project. When the job completes, the withdrawal is committed and the allocation is debited.

- Wright-Patterson aborts a job if it runs over the available balance on the allocation.
- NASA-Ames will allow the overdrafting job to finish, but will not allow any more jobs to be run for that project.
- PSC checks available balances once per day. If a project goes over its allocation, it is reviewed internally before further jobs for that project are denied. This is in support of PSC's philosophy of maximizing machine usage to promote scientific exploration.

Accountability

All sites are accountable to their funding agencies. This accountability takes the form of regular reporting, generally on a monthly or quarterly basis. Each agency, though, has their own way of measuring resource utilization, so each site reports usage differently. For a site like PSC, which receives funding from a number of different agencies, this means that a number of unique reports must be generated, with content and frequency varying appropriately.

In addition to reporting usage to the relevant funding agencies, most sites keep their users informed as to the status of their allocations. NASA-Ames and PNNL send out monthly statements to their PIs; NASA-Ames and PSC also provide tools to let users query their allocation status as needed.

- *What has to be reported?* Because of the diversity of sponsors & funding agencies, each site has different reporting requirements:

NASA-Ames	PSC	PNNL	Wright-Patterson
<ul style="list-style-type: none"> • Machine • Allocation • Usage • Agency (NASA division) • User • Projects 	<ul style="list-style-type: none"> • Machine • Allocation • Usage • Counts: User, Project, State • Agency • Institution • Grant • User 	<ul style="list-style-type: none"> • Nodehours per machine • Availability • Utilization • CPU Efficiency <hr/> <ul style="list-style-type: none"> • Allocation Information (statement from Qbank) • Credits/user 	<ul style="list-style-type: none"> • Project-level utilization • # of users • Queue wait • Wall clock wait • CPU utilization • # of processors used • Expansion factor (by queue for each system or platform) • CPU hours • COTS (package) utilization

- *How is usage calculated?* Usage is calculated differently at each site:

NASA-Ames	PSC	PNNL	Wright-Patterson
CPU + Memory + Wallclock	CPU + Memory + Connect	CPU Only	Wallclock * (# of processors)

Each site calculates usage to measure utilization in standards that are most relevant for their mission. This variety is important because it allows the site to tune their resources to meet specific needs. Differentiation of resources will also be important to GRID users, who will be offered a number of similar platforms and should be able to select those that will best fit their needs.

Conclusions

- Points of commonality
 - All sites support a variety of resources already.
 - All resource requests are reviewed before they are granted.
 - Users must have an account on the specific resource to be permitted to access it.
 - Usage must be reported to the site's sponsoring or funding organizations.
- Points of differentiation
 - Not all requests are reviewed on-site. As a result, sites do not have full autonomy to decide who may access their resources.
 - Review criteria are mission-specific and therefore vary from site to site.
 - The review cycle varies from days to months, even at a given site. Some requests might not be reviewed for up to 6 months; others may be granted access in less than a week.
 - Not all sites accept requests from commercial organizations.
 - "Usage" is calculated differently at all sites.
 - All sites handle overdrafts differently.
 - Sites are accountable to their sponsors, but on different reporting cycles.
 - Different site sponsors require different utilization measures from each sites.
 - Sites have different methods for keeping users up to date on their allocation status.

Not all points of commonality are supportive of GRID computing, but neither are all points of differentiation necessarily barriers. One of the biggest challenges will be addressing the common requirement that users have local accounts to use resources. Local accounts are the basis for usage accounting, which is the primary feedback mechanism between the sites and their funding agencies. On the other hand, variations on how sites review request or calculate usage may allow the sites to maintain their unique identities in the GRID environment, enabling users to make more informed decisions when requesting GRID resources.

This survey was conducted so that the Distributed Accounting Working Group would have a more solid context within which to work. Although the sites that responded are generally small, they are a diverse cross-section of the high-performance computing world. The DAWG will be well-served as it works on recommendations and standards to keep in mind the needs of these sites.

Appendix A

Survey Questions

QUESTIONS ABOUT YOUR SITE:

- Which site are you?
- What resources do you have available?
- How many current users does your site support?
- How many active grants/projects does your site support?
- How many staff members perform allocations activities as their primary or as a major job function?
- How many staff members perform usage accounting activities as their primary or as a major job function?
- How much overlap is there between allocations and usage accounting staff (i.e. do the same people do both sets of tasks, or are the tasks divided between parts of your organization)?

CURRENT REPORTING REQUIREMENTS:

- To what outside entities are you accountable?
- How often must you report to them?
- What must you report?

GENERAL:

- Describe the allocations process at your site.
- Describe the resource usage reporting process at your site.

USERS:

- Does your site require a local account for a user to utilize your resources?
- At what level are resources allocated; e.g. project (a group of users who share an allocation) vs. user (individual allocations for individual people)?
- Are there any differences between how cash (e.g. commercial or industrial) vs. sponsored (e.g. NSF, NIH) requests are granted or reported?

RESOURCE REQUEST REVIEWS:

- Are requests for resources or allocations reviewed before they are granted?
- If so, what are the criteria for review (e.g. scientific merit, user expertise, etc)?
- How long does the review process take, from submission of the request until it is granted or denied?

RESOURCES:

- Do you use atomic (e.g. CPU hours only) or composite (e.g. CPU hours + disk used) measures to track resource usage?
- If you use composites, what components are tracked?
- Do the components carry equal weight? If not, how are weights determined?
- If your site has multiple resources available, are allocations made for the individual resources, or are they made across all platforms at the site?
- How often is usage data collected, analyzed, and reported?
- Are resources at your site available for interactive use, or only via batch?
- Are there differences in allocations and/or usage accounting & reporting depending on the mode of access?
- If a job, user, or project exceeds its requested/allocated resource limits, what happens (e.g. abend/abort, finish run & shut off, etc)?